

# Local convergence analysis of inexact Gauss-Newton like methods under majorant condition

O. P. Ferreira\*      M. L. N. Gonçalves<sup>†</sup>      P. R. Oliveira<sup>‡</sup>

August 12, 2010

## Abstract

In this paper, we present a local convergence analysis of inexact Gauss-Newton like methods for solving nonlinear least squares problems. Under the hypothesis that the derivative of the function associated with the least square problem satisfies a majorant condition, we obtain that the method is well-defined and converges. Our analysis provides a clear relationship between the majorant function and the function associated with the least square problem. It also allows us to obtain an estimate of convergence ball for inexact Gauss-Newton like methods and some important, special cases.

**Keywords:** Nonlinear least squares problems; inexact Gauss-Newton like methods; Majorant condition; Local convergence.

## 1 Introduction

Let  $\mathbb{X}$  and  $\mathbb{Y}$  be real or complex Hilbert spaces. Let  $\Omega \subseteq \mathbb{X}$  be an open set, and  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable nonlinear function. Consider the following *nonlinear least squares* problems

$$\min_{x \in \Omega} \|F(x)\|^2. \quad (1)$$

The interest in this problem arises in data fitting, when  $\mathbb{X} = \mathbb{R}^n$  and  $\mathbb{Y} = \mathbb{R}^m$  and  $m$  is the number of observations and  $n$  is the number of parameters, see for example [13]. A solution  $x_* \in \Omega$  of (1) is also called a least-squares solution of nonlinear equation  $F(x) = 0$ .

---

\*IME/UFG, Campus II- Caixa Postal 131, 74001-970 - Goiânia, GO, Brazil (E-mail:orizon@mat.ufg.br). The author was partly supported by CNPq Grant 473756/2009-9, CNPq Grant 302024/2008-5, PRONEX-Optimization(FAPERJ/CNPq) and FUNAPE/UFG.

<sup>†</sup>COPPE-Sistemas, Universidade Federal do Rio de Janeiro, 21945-970 Rio de Janeiro, RJ, BR (E-mail:maxlmg@cos.ufrj.br). The author was partly supported by CNPq Grant 473756/2009-9.

<sup>‡</sup>COPPE-Sistemas, Universidade Federal do Rio de Janeiro, 21945-970 Rio de Janeiro, RJ, BR (Email:poliveir@cos.ufrj.br). This author was partly supported by CNPq.

When  $F'(x)$  is injective and has closed image for all  $x \in \Omega$ , the Gauss-Newton's method finds stationary points of the above problem. Formally, the Gauss-Newton's method is described as follows: Given an initial point  $x_0 \in \Omega$ , define

$$x_{k+1} = x_k + S_k, \quad F'(x_k)^* F'(x_k) S_k = -F'(x_k)^* F(x_k), \quad k = 0, 1, \dots,$$

where  $A^*$  denotes the adjoint of the operator  $A$ . It is worth pointing out that if  $x_*$  is solution of (1),  $F(x_*) = 0$  and  $F'(x_*)$  is invertible, then the theories of the Gauss-Newton's method merge into the theories of Newton's method. Early works dealing with the convergence of Newton's and Gauss-Newton's methods include [1, 2, 3, 6, 8, 9, 10, 11, 14, 16, 17, 19, 20, 21, 24, 25, 28].

The inexact Gauss-Newton process is described as follows: Given an initial point  $x_0 \in \Omega$ , define

$$x_{k+1} = x_k + S_k, \quad k = 0, 1, \dots,$$

where  $B_k : \mathbb{X} \rightarrow \mathbb{Y}$  is a linear operator and  $S_k$  is any approximated solution of the linear system

$$B_k S_k = -F'(x_k)^* F(x_k) + r_k,$$

for a suitable residual  $r_k \in \mathbb{Y}$ . In particular, the above process is *inexact Gauss-Newton method* if  $B_k = F'(x_k)^T F'(x_k)$ , the process is *inexact modified Gauss-Newton method* if  $B_k = F'(x_0)^T F'(x_0)$ , and it represents a *inexact Gauss-Newton like method* if  $B_k$  is an approximation of  $F'(x_k)^T F'(x_k)$ .

For inexact Newton methods, as shown in [12], if  $\|r_k\| \leq \theta_k \|F(x_k)\|$  for  $k = 0, 1, \dots$  and  $\{\theta_k\}$  is a sequence of forcing terms such that  $0 \leq \theta_k < 1$  then there exists  $\epsilon > 0$  such that the sequence  $\{x_k\}$ , for any initial point  $x_0 \in B(x_*, \epsilon) = \{x \in \mathbb{R}^n : \|x_* - x\| < \epsilon\}$ , is well defined and converges linearly to  $x_*$  in the norm  $\|y\|_* = \|F'(x_*)y\|$ , where  $\|\cdot\|$  is any norm in  $\mathbb{R}^n$ . As pointed out by [22] (see also [23]) the result of [12] is difficult to apply due to a dependence of the norm  $\|\cdot\|_*$ , which is not computable.

Formally, the inexact Gauss-Newton like methods for solving (1), which we will consider, are described as follows: Given an initial point  $x_0 \in \Omega$ , define

$$x_{k+1} = x_k + S_k, \quad B(x_k) S_k = -F'(x_k)^* F(x_k) + r_k, \quad k = 0, 1, \dots,$$

where  $B(x_k)$  is a suitable invertible approximation of the derivative  $F'(x_k)^* F'(x_k)$  and the residual tolerance  $r_k$  and the preconditioning invertible matrix  $P_k$  (considered for the first time in [23]) for the linear system defining the step  $S_k$  satisfy

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^* F(x_k)\|,$$

for suitable forcing number  $\theta_k$ . Note that, if the forcing sequence vanishes, i.e.,  $\theta_k = 0$  for all  $k$ , the inexact Gauss-Newton methods include the class of Gauss-Newton iterative methods. Hence, the theories of inexact Gauss-Newton methods merge into the theories of Gauss-Newton methods.

The classical local convergence analysis for the inexact Newton's methods (see [12, 23]) requires, among other hypotheses, that  $F'$  satisfies the Lipschitz condition. In the last years, there have

been papers dealing with the issue of convergence of the Newton method and inexact Newton's method, including the Gauss-Newton's method and inexact Gauss-Newton's method, by relaxing the assumption of Lipschitz continuity of the derivative (see for example: [5, 6, 7, 9, 14, 15, 16, 17, 21, 28]). One of the main conditions that relaxes the condition of the Lipschitz continuity of the derivative is the majorant condition, which we will use, and Wang's condition, introduced in [28] and used in [5, 6, 7, 8, 20, 21] to study the Gauss-Newton's and Newton's methods. In fact, it can be shown that these conditions are equivalent. But the formulation as a majorant condition is in some sense better than Wang's condition, as it provides a clear relationship between the majorant function and the nonlinear function under consideration. Besides, the majorant condition provides a simpler proof of convergence.

In the present paper, we are interested in the local convergence analysis, i.e., based on the information in a neighbourhood of a stationary point of (1) we determine the convergence ball of the method. Following the ideas of [14, 15, 16, 17], we will present a new local convergence analysis for inexact Gauss-Newton like methods under majorant condition. The convergence analysis presented provides a clear relationship between the majorant function, which relaxes the Lipschitz continuity of the derivate, and the function associated with the nonlinear least square problem (see for example: Lemmas 12, 13 and 14). Besides, the results presented here have the conditions and the proof of convergence in quite a simple manner. Moreover, two unrelated previous results pertaining to inexact Gauss-Newton like methods are unified, namely, the result for analytical functions and the classical one for functions with Lipschitz derivative.

The organization of the paper is as follows. In Section 1.1, we list some notations and basic results used in our presentation. In Section 2 the main result is stated, and in Section 2.1 some properties involving the majorant function are established. In Section 2.2 we present the relationships between the majorant function and the non-linear function  $F$ . In Section 2.3 the main result is proven and some applications of this result are given in Section 3. Some final remarks are offered in Section 4.

## 1.1 Notation and auxiliary results

The following notations and results are used throughout our presentation. Let  $\mathbb{X}$  and  $\mathbb{Y}$  be Hilbert spaces. The open and closed ball at  $a \in \mathbb{X}$  and radius  $\delta > 0$  are denoted, respectively by

$$B(a, \delta) := \{x \in \mathbb{X}; \|x - a\| < \delta\}, \quad B[a, \delta] := \{x \in \mathbb{X}; \|x - a\| \leq \delta\}.$$

The set  $\Omega \subseteq \mathbb{X}$  is an open set and the function  $F : \Omega \rightarrow \mathbb{Y}$  is continuously differentiable, and  $F'(x)$  has closed image in  $\Omega$ .

Let  $A : \mathbb{X} \rightarrow \mathbb{Y}$  be a continuous and injective linear operator with closed image. The Moore-Penrose inverse  $A^\dagger : \mathbb{Y} \rightarrow \mathbb{X}$  of  $A$  is defined by

$$A^\dagger := (A^* A)^{-1} A^*,$$

where  $A^*$  denotes the adjoint of the linear operator  $A$ .

**Lemma 1.** (*Banach's Lemma*) Let  $B : \mathbb{X} \rightarrow \mathbb{X}$  be a continuous linear operator, and  $I : \mathbb{X} \rightarrow \mathbb{X}$  the identity operator. If  $\|B - I\| < 1$ , then  $B$  is invertible and  $\|B^{-1}\| \leq 1/(1 - \|B - I\|)$ .

*Proof.* See the proof of Lemma 1, p. 189 of Smale [26] with  $A = I$  and  $c = \|B - I\|$ .  $\square$

**Lemma 2.** Let  $A, B : \mathbb{X} \rightarrow \mathbb{Y}$  be a continuous linear operator with closed image. If  $A$  is injective,  $E = B - A$  and  $\|EA^\dagger\| < 1$ , then  $B$  is injective.

*Proof.* In fact,  $B = A + E = (I + EA^\dagger)A$ , from the condition  $\|EA^\dagger\| < 1$ , we have of Lemma 1 that  $I + EA^\dagger$  is invertible. So,  $B$  is injective.  $\square$

The next lemma is proven in Stewart [27] ( see also, Wedin [29] ) for  $m \times n$  matrix with  $m \geq n$  and  $\text{rank}(A) = \text{rank}(B) = n$ , that proof holds in more general context as we will state below.

**Lemma 3.** Let  $A, B : \mathbb{X} \rightarrow \mathbb{Y}$  be continuous and injective linear operators with closed images. Assume that  $E = B - A$  and  $\|A^\dagger\|\|E\| < 1$ , then

$$\|B^\dagger\| \leq \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|E\|}, \quad \|B^\dagger - A^\dagger\| \leq \frac{\sqrt{2}\|A^\dagger\|^2\|E\|}{1 - \|A^\dagger\|\|E\|}.$$

**Proposition 4.** If  $0 \leq t < 1$ , then  $\sum_{i=0}^{\infty} (i+2)(i+1)t^i = 2/(1-t)^3$ .

*Proof.* Take  $k = 2$  in Lemma 3, pp. 161 of Blum, et al. [4].  $\square$

Also, the following auxiliary results of elementary convex analysis will be needed:

**Proposition 5.** Let  $R > 0$ . If  $\varphi : [0, R) \rightarrow \mathbb{R}$  is convex, then

$$D^+\varphi(0) = \lim_{u \rightarrow 0+} \frac{\varphi(u) - \varphi(0)}{u} = \inf_{0 < u} \frac{\varphi(u) - \varphi(0)}{u}.$$

*Proof.* See Theorem 4.1.1 on pp. 21 of Hiriart-Urruty and Lemaréchal [18].  $\square$

**Proposition 6.** Let  $\epsilon > 0$  and  $\tau \in [0, 1]$ . If  $\varphi : [0, \epsilon) \rightarrow \mathbb{R}$  is convex, then  $l : (0, \epsilon) \rightarrow \mathbb{R}$  defined by

$$l(t) = \frac{\varphi(t) - \varphi(\tau t)}{t},$$

is increasing.

*Proof.* See Theorem 4.1.1 and Remark 4.1.2 on pp. 21 of Hiriart-Urruty and Lemaréchal [18].  $\square$

## 2 Local analysis for inexact Gauss-Newton like methods

In this section, we will state and prove a local theorem for inexact Gauss-Newton like methods. Assuming that the function

$$\Omega \ni x \mapsto F(x)^* F(x),$$

has a point stationary  $x_*$ , we will, under mild conditions, prove that the inexact Gauss-Newton like methods is well defined and that the generated sequence converges linearly to this point stationary. The statement of the theorem is as follows:

**Theorem 7.** *Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and*

$$c := \|F(x_*)\|, \quad \beta := \|F'(x_*)^\dagger\|, \quad \kappa := \sup \{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

*Suppose that  $F'(x_*)^* F(x_*) = 0$ ,  $F'(x_*)$  is injective and there exists a  $f : [0, R) \rightarrow \mathbb{R}$  continuously differentiable such that*

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq f'(\|x - x_*\|) - f'(\tau\|x - x_*\|), \quad (2)$$

*for all  $\tau \in [0, 1]$ ,  $x \in B(x_*, \kappa)$  and*

**h1)**  $f(0) = 0$  and  $f'(0) = -1$ ;

**h2)**  $f'$  is convex and strictly increasing;

**h3)**  $\alpha := \sqrt{2} c \beta^2 D^+ f'(0) < 1$ .

*Take  $0 \leq \vartheta < 1$ ,  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1(\alpha + \alpha\vartheta + \vartheta) + \omega_2 < 1$ . Let the positive constants*

$$\nu := \sup \{t \in [0, R) : \beta[f'(t) + 1] < 1\},$$

$$\rho := \sup \left\{ t \in (0, \nu) : (1 + \vartheta)\omega_1\beta \frac{tf'(t) - f(t) + \sqrt{2}c\beta[f'(t) + 1]}{t[1 - \beta(f'(t) + 1)]} + \omega_1\vartheta + \omega_2 < 1 \right\}, \quad r := \min \{\kappa, \rho\}.$$

*Then, the inexact Gauss-Newton like methods for solving (1), with initial point  $x_0 \in B(x_*, r) \setminus \{x_*\}$*

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F'(x_k)^* F(x_k) + r_k, \quad k = 0, 1, \dots, \quad (3)$$

*for the forcing term  $\theta_k$  and the following conditions for the residual  $r_k$  and the invertible matrix  $P_k$  preconditioning the linear system in (3)*

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^* F(x_k)\|, \quad 0 \leq \theta_k \text{cond}(P_k F'(x_k)^* F'(x_k)) \leq \vartheta, \quad k = 0, 1, \dots,$$

where  $B(x_k)$  is an invertible approximation of  $F'(x_k)^*F'(x_k)$  satisfying the following conditions

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

is well defined, contained in  $B(x_*, r)$ , converges to  $x_*$  and there holds

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq (1 + \vartheta)\omega_1\beta \frac{[f'(\|x_0 - x_*\|)\|x_0 - x_*\| - f(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f'(\|x_0 - x_*\|) + 1)]} \|x_k - x_*\|^2 \\ &+ \left( \frac{(1 + \vartheta)\omega_1\sqrt{2}c\beta^2[f'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f'(\|x_0 - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2 \right) \|x_k - x_*\|, \quad k = 0, 1, \dots \end{aligned} \quad (4)$$

**Remark 1.** In particular, if taking  $\vartheta = 0$  (in this case  $\theta_k \equiv 0$  and  $r_k \equiv 0$ ) in Theorem 7, we obtain the convergence of Gauss-Newton's like method under majorant condition which, for  $\omega_1 = 1$  and  $\omega_2 = 0$ , i.e.,  $B(x_k) = F'(x_k)^*F'(x_k)$ , has been obtained by Ferreira et al. [16] in Theorem 7. Now, if taking  $c = 0$  (the so-called zero-residual case) and  $F'(x_*)$  is invertible, we obtain the convergence of inexact Newton-Like methods under majorant condition, which has been obtained by Ferreira, Gonçalves [15] in Theorem 4. Finally, if  $c = \vartheta = \omega_2 = 0$ ,  $\omega_1 = 1$  and  $F'(x_*)$  is invertible in Theorem 7, we obtain the convergence of Newton's method under majorant condition, which has been obtained by Ferreira [14] in Theorem 2.1.

For the important case  $\vartheta = 0$ , namely, Gauss-Newton's like method under majorant condition, the Theorem 7 becomes:

**Corollary 8.** Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and

$$c := \|F(x_*)\|, \quad \beta := \|F'(x_*)^\dagger\|, \quad \kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

Suppose that  $F'(x_*)^*F'(x_*) = 0$ ,  $F'(x_*)$  is injective and there exists a  $f : [0, R) \rightarrow \mathbb{R}$  continuously differentiable such that

$$\|F'(x) - F'(x_* + \tau(x - x_*))\| \leq f'(\|x - x_*\|) - f'(\tau\|x - x_*\|),$$

for all  $\tau \in [0, 1]$ ,  $x \in B(x_*, \kappa)$  and

**h1)**  $f(0) = 0$  and  $f'(0) = -1$ ;

**h2)**  $f'$  is convex and strictly increasing;

**h3)**  $\alpha := \sqrt{2}c\beta^2D^+f'(0) < 1$ .

Take  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1\alpha + \omega_2 < 1$ . Let  $\nu := \sup \{t \in [0, R) : \beta[f'(t) + 1] < 1\}$ ,

$$\rho := \sup \left\{ t \in (0, \nu) : \omega_1 \beta \frac{tf'(t) - f(t) + \sqrt{2}c\beta[f'(t) + 1]}{t[1 - \beta(f'(t) + 1)]} + \omega_2 < 1 \right\}, \quad r := \min \{\kappa, \rho\}.$$

Then, the Gauss-Newton's like method for solving (1), with initial point  $x_0 \in B(x_*, r) \setminus \{x_*\}$

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F'(x_k)^*F(x_k), \quad k = 0, 1, \dots,$$

where  $B(x_k)$  is an invertible approximation of  $F'(x_k)^*F'(x_k)$  satisfying

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

is well defined, contained in  $B(x_*, r)$ , converges to  $x_*$  and there holds

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \omega_1 \beta \frac{[f'(\|x_0 - x_*\|)\|x_0 - x_*\| - f(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f'(\|x_0 - x_*\|) + 1)]} \|x_k - x_*\|^2 \\ &\quad + \left( \frac{\omega_1 \sqrt{2}c\beta^2[f'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f'(\|x_0 - x_*\|) + 1)]} + \omega_2 \right) \|x_k - x_*\|, \quad k = 0, 1, \dots \end{aligned} \quad (5)$$

**Remark 2.** Despite the fact that the above corollary is a special case of Theorem 7, the results contained therein extend the results of Chen and Li in [8], as the results obtained [8] are only for the case  $c = 0$ .

**Remark 3.** Assumption (2) is crucial for our analysis. It should be pointed that, under appropriate regularity conditions in the nonlinear function  $F$ , assumption (2) always holds on a suitable neighbourhood of  $x_*$ . For instance, if  $F$  is two times continuously differentiable, then the majorant function  $f : [0, \kappa] \rightarrow \mathbb{R}$ , as defined by  $f(t) = Kt^2/2 - t$ , where  $K = \sup\{\|F''(x)\| : x \in B[x_*, \kappa]\}$  satisfies assumption (2). Estimating the constant  $K$  is a very difficult problem. Therefore, the goal is to identify classes of nonlinear functions for which it is possible to obtain a majorant function. We will give some examples of such classes in Section 3.

To prove Theorem 7 we need some results. From here on, we assume that all assumptions of Theorem 7 hold.

## 2.1 The majorant function

In this section, we will prove that the constant  $\kappa$  associated with  $\Omega$  and the constants  $\nu$ ,  $\rho$  and  $r$  associated with the majorant function  $f$  are positive. We will also prove some results related to the function  $f$ .

We begin by noting that  $\kappa > 0$ , because  $\Omega$  is an open set and  $x_* \in \Omega$ .

**Proposition 9.** *The constant  $\nu$  is positive and there holds*

$$\beta[f'(t) + 1] < 1, \quad t \in (0, \nu).$$

*Proof.* As  $f'$  is continuous in  $(0, R)$  and  $f'(0) = -1$ , it is easy to conclude that

$$\lim_{t \rightarrow 0} \beta[f'(t) + 1] = 0.$$

Thus, there exists a  $\delta > 0$  such that  $\beta(f'(t) + 1) < 1$  for all  $t \in (0, \delta)$ . Hence,  $\nu > 0$ .

Using **h2** and definition of  $\nu$  the last part of the proposition follows. □

**Proposition 10.** *The following functions are increasing:*

- i)  $[0, R) \ni t \mapsto 1/[1 - \beta(f'(t) + 1)]$ ;
- ii)  $(0, R) \ni t \mapsto [tf'(t) - f(t)]/t^2$ ;
- iii)  $(0, R) \ni t \mapsto [f'(t) + 1]/t$ ;

*As a consequence, there is an increase of the following functions*

$$(0, R) \ni t \mapsto \frac{tf'(t) - f(t)}{t^2[1 - \beta(f'(t) + 1)]}, \quad (0, R) \ni t \mapsto \frac{f'(t) + 1}{t[1 - \beta(f'(t) + 1)]}.$$

*Proof.* The item **i** is immediate, because  $f'$  is strictly increasing in  $[0, R)$ .

For proving item **ii**, note that after some simple algebraic manipulations we have

$$\frac{tf'(t) - f(t)}{t^2} = \int_0^1 \frac{f'(t) - f'(\tau t)}{t} d\tau.$$

So, applying Proposition 6 with  $f' = \varphi$  and  $\epsilon = R$  the statement follows.

For establishing item **iii** use **h2**,  $f'(0) = -1$  and Proposition 6 with  $f' = \varphi$ ,  $\epsilon = R$  and  $\tau = 0$ .

To prove that the functions in the last part are increasing, combine item **i** with **ii** for the first function, and **i** with **iii** for the second function. □

**Proposition 11.** *The constant  $\rho$  is positive and there holds*

$$(1 + \vartheta)\omega_1\beta \frac{tf'(t) - f(t) + \sqrt{2}c\beta[f'(t) + 1]}{t[1 - \beta(f'(t) + 1)]} + \omega_1\vartheta + \omega_2 < 1, \quad \forall t \in (0, \rho).$$



*Proof.* First of all, note that the assumption **h1** implies, after simple calculation, that

$$\lim_{t \rightarrow 0} \frac{tf'(t) - f(t)}{t[1 - \beta(f'(t) + 1)]} = \lim_{t \rightarrow 0} \frac{f'(t) - (f(t) - f(0))/t}{1 - \beta(f'(t) + 1)} = 0.$$

Again, using **h1**, some algebraic manipulation and that  $f'$  is convex, we have by Proposition 5

$$\lim_{t \rightarrow 0} \frac{f'(t) + 1}{t[1 - \beta(f'(t) + 1)]} = \lim_{t \rightarrow 0} \frac{(f'(t) - f'(0))/t}{1 - \beta(f'(t) + 1)} = D^+ f'(0).$$

Hence, by combining the two above equalities it is easy to conclude that

$$\lim_{t \rightarrow 0} (1 + \vartheta)\omega_1\beta \frac{tf'(t) - f(t) + \sqrt{2}c\beta[f'(t) + 1]}{t[1 - \beta(f'(t) + 1)]} + \omega_1\vartheta + \omega_2 = (1 + \vartheta)\omega_1\sqrt{2}c\beta^2 D^+ f'(0) + \omega_1\vartheta + \omega_2.$$

As,  $\alpha = \sqrt{2}c\beta^2 D^+ f'(0)$  and  $\omega_1(\alpha + \alpha\vartheta + \vartheta) + \omega_2 < 1$ , we obtain that there exists a  $\delta > 0$  such that

$$(1 + \vartheta)\omega_1\beta \frac{tf'(t) - f(t) + \sqrt{2}c\beta[f'(t) + 1]}{t[1 - \beta(f'(t) + 1)]} + \omega_1\vartheta + \omega_2 < 1, \quad t \in (0, \delta),$$

Hence,  $\delta \leq \rho$ , which proves the first statement. To conclude the proof, we use the definition of  $\rho$ , the above inequality, and the last part of Proposition 10.  $\square$

## 2.2 Relationship of the majorant function with the non-linear function

In this section we will present the main relationships between the majorant function  $f$  and the function  $F$  associated with the nonlinear least square problem.

**Lemma 12.** *Let  $x \in \Omega$ . If  $\|x - x_*\| < \min\{\nu, \kappa\}$ , then  $F'(x)^* F'(x)$  is invertible and the following inequalities hold*

$$\|F'(x)^\dagger\| \leq \frac{\beta}{1 - \beta[f'(\|x - x_*\|) + 1]}, \quad \|F'(x)^\dagger - F'(x_*)^\dagger\| < \frac{\sqrt{2}\beta^2[f'(\|x - x_*\|) + 1]}{1 - \beta[f'(\|x - x_*\|) + 1]}.$$

*In particular,  $F'(x)^* F'(x)$  is invertible in  $B(x_*, r)$ .*

*Proof.* Let  $x \in \Omega$  such that  $\|x - x_*\| < \min\{\nu, \kappa\}$ . Since  $\|x - x_*\| < \nu$ , using the definition of  $\beta$ , the inequality (2) and last part of Proposition 9 we have

$$\|F'(x) - F'(x_*)\| \|F'(x_*)^\dagger\| \leq \beta[f'(\|x - x_*\|) - f'(0)] < 1.$$

For the sake of simplicity, the notations define the following matrices

$$A = F'(x_*), \quad B = F'(x), \quad E = F'(x) - F'(x_*). \quad (6)$$

The last definitions, together with the latter inequality, imply that

$$\|EA^\dagger\| \leq \|E\|\|A^\dagger\| < 1,$$

which, using that  $F'(x_*)$  is injective, implies in view of Lemma 2 that  $F'(x)$  is injective. So,  $F'(x)^*F'(x)$  is invertible and by definition of  $r$  we obtain that  $F'(x)^*F'(x)$  is invertible for all  $x \in B(x_*, r)$ .

We already know that  $F'(x_*)$  and  $F'(x)$  are injective. Hence, to conclude the lemma use definitions in (6) and then combine the above inequality and Lemma 3.  $\square$

Now, it is convenient to study the linearization error of  $F$  at point in  $\Omega$ , for which we define

$$E_F(x, y) := F(y) - [F(x) + F'(x)(y - x)], \quad y, x \in \Omega. \quad (7)$$

We will bound this error by the error in the linearization on the majorant function  $f$

$$e_f(t, u) := f(u) - [f(t) + f'(t)(u - t)], \quad t, u \in [0, R]. \quad (8)$$

**Lemma 13.** *If  $\|x - x_*\| < \kappa$ , then there holds  $\|E_F(x, x_*)\| \leq e_f(\|x - x_*\|, 0)$ .*

*Proof.* Since  $B(x_*, \kappa)$  is convex, we obtain that  $x_* + \tau(x - x_*) \in B(x_*, \kappa)$ , for  $0 \leq \tau \leq 1$ . Thus, as  $F$  is continuously differentiable in  $\Omega$ , definition of  $E_F$  and some simple manipulations yield

$$\|E_F(x, x_*)\| \leq \int_0^1 \|[F'(x) - F'(x_* + \tau(x - x_*))]\| \|x_* - x\| d\tau.$$

From the last inequality and the assumption (2), we obtain

$$\|E_F(x, x_*)\| \leq \int_0^1 [f'(\|x - x_*\|) - f'(\tau\|x - x_*\|)] \|x - x_*\| d\tau.$$

Evaluating the above integral and using definition of  $e_f$ , the statement follows.  $\square$

Define the Gauss-Newton step to the functions  $F$  by the following equality:

$$S_F(x) := -F'(x)^\dagger F(x). \quad (9)$$

**Lemma 14.** *If  $\|x - x_*\| < \min\{\nu, \kappa\}$ , then*

$$\|S_F(x)\| \leq \frac{\beta e_f(\|x - x_*\|, 0) + \sqrt{2}c\beta^2[f'(\|x - x_*\|) + 1]}{1 - \beta[f'(\|x - x_*\|) + 1]} + \|x - x_*\|.$$

*Proof.* Using (9),  $F'(x_*)^*F(x_*) = 0$  and some algebraic manipulation, it follows from (7) that

$$\begin{aligned}\|S_F(x)\| &= \|F'(x)^\dagger (F(x_*) - [F(x) + F'(x)(x_* - x)]) - (F'(x)^\dagger - F'(x_*)^\dagger)F(x_*) + (x_* - x)\| \\ &\leq \|F'(x)^\dagger\| \|E_F(x, x_*)\| + \|F'(x)^\dagger - F'(x_*)^\dagger\| \|F(x_*)\| + \|x - x_*\|.\end{aligned}$$

So, the last inequality together with the Lemma 12, Lemma 13 and definition of  $c$ , imply that

$$\|S_F(x)\| \leq \frac{\beta e_f(\|x - x_*\|, 0)}{1 - \beta[f'(\|x - x_*\|) + 1]} + \frac{\sqrt{2}c\beta^2[f'(\|x - x_*\|) + 1]}{1 - \beta[f'(\|x - x_*\|) + 1]} + \|x - x_*\|,$$

which is equivalent to the desired inequality.  $\square$

**Lemma 15.** *Let  $\Omega \subseteq \mathbb{X}$  be an open set and  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and  $c, \beta, \kappa$  as a definition in Theorem 7. Suppose that  $F'(x_*)^*F(x_*) = 0$ ,  $F'(x_*)$  is injective and there exists a  $f : [0, R) \rightarrow \mathbb{R}$  continuously differentiable satisfying (2), **h1**, **h2** and **h3**. Let  $\alpha, \vartheta, \omega_1, \omega_2, \nu, \rho$  and  $r$  as in Theorem 7. Assume that  $x \in B(x_*, r) \setminus \{x_*\}$ , i.e.,  $0 < \|x - x_*\| < r$ . Define*

$$x_+ = x + S, \quad B(x)S = -F'(x)^*F(x) + r, \quad (10)$$

where  $B(x)$  is an invertible approximation of  $F'(x)^*F'(x)$  satisfying

$$\|B(x)^{-1}F'(x)^*F'(x)\| \leq \omega_1, \quad \|B(x)^{-1}F'(x)^*F'(x) - I\| \leq \omega_2, \quad (11)$$

and the forcing term  $\theta$  and the residual  $r$  satisfy

$$\theta \text{cond}(PF'(x)^*F'(x)) \leq \vartheta, \quad \|Pr\| \leq \theta \|PF'(x)^*F(x)\|, \quad (12)$$

with  $P$  an invertible matrix (preconditioner for the linear system in (10)). Then  $x_+$  is well defined and there holds

$$\begin{aligned}\|x_+ - x_*\| &\leq (1 + \vartheta)\omega_1\beta \frac{[f'(\|x - x_*\|)\|x - x_*\| - f(\|x - x_*\|)]}{\|x - x_*\|^2[1 - \beta(f'(\|x - x_*\|) + 1)]} \|x - x_*\|^2 \\ &\quad + \left( \frac{(1 + \vartheta)\omega_1\sqrt{2}c\beta^2[f'(\|x - x_*\|) + 1]}{\|x - x_*\|[1 - \beta(f'(\|x - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2 \right) \|x - x_*\|, \quad k = 0, 1, \dots \quad (13)\end{aligned}$$

In particular,

$$\|x_+ - x_*\| < \|x - x_*\|.$$

*Proof.* First note that, as  $\|x - x_*\| < r$ , it follows from Lemma 12 that  $F'(x)^*F'(x)$  is invertible. Now, let  $B(x)$  an invertible approximation of it satisfying (11). Thus,  $x_+$  is well defined. Now, as  $F'(x_*)^*F(x_*) = 0$ , some simple algebraic manipulation and (10) yield

$$\begin{aligned}x_+ - x_* &= x - x_* - B(x)^{-1}F'(x)^*(F(x) - F(x_*)) + B(x)^{-1}r \\ &\quad + B(x)^{-1}F'(x)^*F'(x) \left[ F'(x_*)^\dagger F(x_*) - F'(x)^\dagger F(x_*) \right].\end{aligned}$$

Again, some algebraic manipulation in the above equation gives

$$\begin{aligned} x_+ - x_* &= B(x)^{-1} F'(x)^* F'(x) F'(x)^\dagger (F(x_*) - [F(x) + F'(x)(x_* - x)]) + B(x)^{-1} r \\ &\quad + B(x)^{-1} (F'(x)^* F'(x) - B(x)) (x - x_*) + B(x)^{-1} F'(x)^* F'(x) [F'(x_*)^\dagger F(x_*) - F'(x)^\dagger F(x_*)]. \end{aligned}$$

The last equation, together with (7) and (11), imply that

$$\|x_+ - x_*\| \leq \omega_1 \|F'(x)^\dagger\| \|E_F(x, x_*)\| + \|B(x)^{-1} r\| + \omega_2 \|x - x_*\| + \omega_1 \|F'(x)^\dagger - F'(x_*)^\dagger\| \|F(x_*)\|.$$

On the other hand, using (9), (11) and (12) we have, by simple calculus,

$$\begin{aligned} \|B(x)^{-1} r\| &\leq \|B(x)^{-1} P^{-1}\| \|Pr\| \\ &\leq \theta \|B(x)^{-1} F'(x)^* F'(x)\| \|(PF'(x)^* F'(x))^{-1}\| \|PF'(x)^* F'(x)\| \|F'(x)^\dagger F(x)\| \\ &\leq \omega_1 \vartheta \|S_F(x)\|. \end{aligned}$$

Hence, it follows from the two last equations that

$$\|x_+ - x_*\| \leq \omega_1 \|F'(x)^\dagger\| \|E_F(x, x_*)\| + \omega_1 \vartheta \|S_F(x)\| + \omega_2 \|x - x_*\| + \omega_1 \|F'(x)^\dagger - F'(x_*)^\dagger\| \|F(x_*)\|.$$

Combining the last equation with the Lemmas 12, 13 and 14, we obtain that

$$\|x_+ - x_*\| \leq (1 + \vartheta) \beta \omega_1 \frac{e_f(\|x - x_*\|, 0) + \sqrt{2} c \beta (f'(\|x - x_*\|) + 1)}{1 - \beta (f'(\|x - x_*\|) + 1)} + \omega_1 \vartheta \|x - x_*\| + \omega_2 \|x - x_*\|.$$

Now, using (8) and some algebraic manipulation, we conclude from the last inequality that

$$\begin{aligned} \|x_+ - x_*\| &\leq (1 + \vartheta) \beta \omega_1 \frac{f'(\|x - x_*\|) \|x - x_*\| - f(\|x - x_*\|) + \sqrt{2} c \beta (f'(\|x - x_*\|) + 1)}{1 - \beta (f'(\|x - x_*\|) + 1)} \\ &\quad + \omega_1 \vartheta \|x - x_*\| + \omega_2 \|x - x_*\|, \end{aligned}$$

which is equivalent to (13). To end the proof, note that the right hand side of (13) is equivalent to

$$\left[ (1 + \vartheta) \omega_1 \beta \frac{f'(\|x - x_*\|) \|x - x_*\| - f(\|x - x_*\|) + \sqrt{2} c \beta (f'(\|x - x_*\|) + 1)}{\|x - x_*\| [1 - \beta (f'(\|x - x_*\|) + 1)]} + \omega_1 \vartheta + \omega_2 \right] \|x - x_*\|.$$

On the other hand, as  $x \in B(x_*, r) \setminus \{x_*\}$ , i.e.,  $0 < \|x - x_*\| < r \leq \rho$  we apply the Proposition 11 with  $t = \|x - x_*\|$  to conclude that the quantity in the bracket above is less than one. So, the last inequality of the lemma follows.  $\square$

### 2.3 Proof of Theorem 7

Now, we will produce the proof of Theorem 7.

*Proof.* Since  $x_0 \in B(x_*, r) \setminus \{x_*\}$ , i.e.,  $0 < \|x_0 - x_*\| < r$ , by combination of Lemma 12, last inequality in Lemma 15 and induction argument, it is easy to see that  $\{x_k\}$  is well defined and remains in  $B(x_*, r)$ .

We are going to prove that  $\{x_k\}$  converges towards  $x_*$ . As,  $\{x_k\}$  is well defined and contained in  $B(x_*, r)$ , applying Lemma 15 with  $x_+ = x_{k+1}$ ,  $x = x_k$ ,  $r = r_k$ ,  $B(x) = B(x_k)$ ,  $P = P_k$ , and  $\theta = \theta_k$  we obtain

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq (1 + \vartheta)\omega_1\beta \frac{[f'(\|x_k - x_*\|)\|x_k - x_*\| - f(\|x_k - x_*\|)]}{\|x_k - x_*\|^2[1 - \beta(f'(\|x_k - x_*\|) + 1)]} \|x_k - x_*\|^2 \\ &\quad + \left( \frac{(1 + \vartheta)\omega_1\sqrt{2}c\beta^2[f'(\|x_k - x_*\|) + 1]}{\|x_k - x_*\|[1 - \beta(f'(\|x_k - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2 \right) \|x_k - x_*\|, \quad k = 0, 1, \dots \end{aligned}$$

Now, using the last inequality of Lemma 15, it is easy to conclude that

$$\|x_k - x_*\| < \|x_0 - x_*\|, \quad k = 1, 2, \dots \quad (14)$$

Hence, combining the last two inequalities with the last part of Proposition 10 we obtain that

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq (1 + \vartheta)\omega_1\beta \frac{[f'(\|x_0 - x_*\|)\|x_0 - x_*\| - f(\|x_0 - x_*\|)]}{\|x_0 - x_*\|^2[1 - \beta(f'(\|x_0 - x_*\|) + 1)]} \|x_k - x_*\|^2 \\ &\quad + \left( \frac{(1 + \vartheta)\omega_1\sqrt{2}c\beta^2[f'(\|x_0 - x_*\|) + 1]}{\|x_0 - x_*\|[1 - \beta(f'(\|x_0 - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2 \right) \|x_k - x_*\|, \quad k = 0, 1, \dots, \end{aligned}$$

which is the inequality (5). Now, using (14) and the last inequality we have

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \\ &\left[ (1 + \vartheta)\omega_1\beta \frac{f'(\|x_0 - x_*\|)\|x_0 - x_*\| - f(\|x_0 - x_*\|) + \sqrt{2}c\beta(f'(\|x_0 - x_*\|) + 1)}{\|x_0 - x_*\|[1 - \beta(f'(\|x_0 - x_*\|) + 1)]} + \omega_1\vartheta + \omega_2 \right] \|x_k - x_*\|, \end{aligned}$$

for all  $k = 0, 1, \dots$ . Applying Proposition 11 with  $t = \|x_0 - x_*\|$  it is straightforward to conclude from the latter inequality that  $\{\|x_k - x_*\|\}$  converges to zero. So,  $\{x_k\}$  converges to  $x_*$ .  $\square$

### 3 Special cases

In this section, we present two special cases of Theorem 7. They include the classical convergence theorem on Gauss-Newton's method under the Lipschitz condition and Smale's theorem on Gauss-Newton for analytical functions.

### 3.1 Convergence result for Lipschitz condition

In this section we show a correspondent theorem for Theorem 7 under the Lipschitz condition, instead of the general assumption (2).

**Theorem 16.** *Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and*

$$c := \|F(x_*)\|, \quad \beta := \|F'(x_*)^\dagger\|, \quad \kappa := \sup \{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

*Suppose that  $F'(x_*)^*F(x_*) = 0$ ,  $F'(x_*)$  is injective and there exists a  $K > 0$  such that*

$$\alpha := \sqrt{2}c\beta^2K < 1, \quad \|F'(x) - F'(y)\| \leq K\|x - y\|, \quad \forall x, y \in B(x_*, \kappa).$$

*Take  $0 \leq \vartheta < 1$ ,  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1(\alpha + \alpha\vartheta + \vartheta) + \omega_2 < 1$ . Let*

$$r := \min \left\{ \kappa, \frac{2(1 - \omega_1\vartheta - \omega_2) - 2\sqrt{2}cK\beta^2\omega_1(1 + \vartheta)}{\beta K(2 + \omega_1 - \vartheta\omega_1 - 2\omega_2)} \right\}.$$

*Then, the inexact Gauss-Newton like methods for solving (1), with initial point  $x_0 \in B(x_*, r) \setminus \{x_*\}$*

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F'(x_k)^*F(x_k) + r_k, \quad k = 0, 1, \dots, \quad (15)$$

*with the following conditions for the residual  $r_k$ , and the forcing term  $\theta_k$*

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^*F(x_k)\|, \quad 0 \leq \theta_k \text{cond}(P_k F'(x_k)^*F'(x_k)) \leq \vartheta, \quad k = 0, 1, \dots,$$

*where  $\{P_k\}$  is an invertible matrix sequence (preconditioners for the linear system in (15)) and  $B(x_k)$  is an invertible approximation of  $F'(x_k)^*F'(x_k)$  satisfying*

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

*is well defined, contained in  $B(x_*, r)$ , converges to  $x_*$  and there holds*

$$\|x_{k+1} - x_*\| \leq \frac{(1 + \vartheta)\beta\omega_1K}{2(1 - \beta K\|x_0 - x_*\|)}\|x_k - x_*\|^2 + \left( \frac{(1 + \vartheta)\omega_1\sqrt{2}c\beta^2K}{1 - \beta K\|x_0 - x_*\|} + \omega_1\vartheta + \omega_2 \right) \|x_k - x_*\|,$$

*for all  $k = 0, 1, \dots$ .*

*Proof.* It is immediate to prove that  $F$ ,  $x_*$  and  $f : [0, \kappa) \rightarrow \mathbb{R}$  as defined by  $f(t) = Kt^2/2 - t$ , satisfy the inequality (2), conditions **h1** and **h2**. Since  $\sqrt{2}c\beta^2K < 1$  the condition **h3** also holds. In this case, it is easy to see that constants  $\nu$  and  $\rho$  as defined in Theorem 7, satisfy

$$0 < \rho = \frac{2(1 - \omega_1\vartheta - \omega_2) - 2\sqrt{2}cK\beta^2\omega_1(1 + \vartheta)}{\beta K(2 + \omega_1 - \vartheta\omega_1 - 2\omega_2)} \leq \nu = 1/\beta K,$$

as a consequence,  $0 < r = \min\{\kappa, \rho\}$ . Therefore, as  $F$ ,  $r$ ,  $f$  and  $x_*$  satisfy all of the hypotheses of Theorem 7, taking  $x_0 \in B(x_*, r) \setminus \{x_*\}$  the statements of the theorem follow from Theorem 7.  $\square$

For the case  $\vartheta = 0$ , the Theorem 16 becomes:

**Corollary 17.** *Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and*

$$c := \|F(x_*)\|, \quad \beta := \left\| F'(x_*)^\dagger \right\|, \quad \kappa := \sup \{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

*Suppose that  $F'(x_*)^* F(x_*) = 0$ ,  $F'(x_*)$  is injective and there exists a  $K > 0$  such that*

$$\alpha := \sqrt{2}c\beta^2 K < 1, \quad \left\| F'(x) - F'(y) \right\| \leq K\|x - y\|, \quad \forall x, y \in B(x_*, \kappa).$$

*Take  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1\alpha + \omega_2 < 1$ . Let*

$$r := \min \left\{ \kappa, \frac{2(1 - \omega_2) - 2\sqrt{2}cK\beta^2\omega_1}{\beta K(2 + \omega_1 - 2\omega_2)} \right\}.$$

*Then, the Gauss-Newton's like method for solving (1), with initial point  $x_0 \in B(x_*, r) \setminus \{x_*\}$*

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F'(x_k)^* F(x_k), \quad k = 0, 1, \dots, \quad (16)$$

*where  $B(x_k)$  is an invertible approximation of  $F'(x_k)^* F'(x_k)$  satisfying*

$$\|B(x_k)^{-1} F'(x_k)^* F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1} F'(x_k)^* F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

*is well defined, contained in  $B(x_*, r)$ , converges to  $x_*$  and there holds*

$$\|x_{k+1} - x_*\| \leq \frac{\beta\omega_1 K}{2(1 - \beta K\|x_0 - x_*\|)} \|x_k - x_*\|^2 + \left( \frac{\omega_1 \sqrt{2}c\beta^2 K}{1 - \beta K\|x_0 - x_*\|} + \omega_2 \right) \|x_k - x_*\|,$$

*for all  $k = 0, 1, \dots$ .*

Note that letting  $c = 0$  in the above corollary, we obtain the Corollary 6.1 of [8].

### 3.2 Convergence result under Smale's condition

In this section we present a correspondent theorem to Theorem 7 under Smale's condition. For more details see Smale [26] and Dedieu and Shub [11].

**Theorem 18.** *Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and*

$$c := \|F(x_*)\|, \quad \beta := \left\| F'(x_*)^\dagger \right\|, \quad \kappa := \sup \{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

Suppose that  $F'(x_*)^*F(x_*) = 0$ ,  $F'(x_*)$  is injective and

$$\gamma := \sup_{n \geq 1} \left\| \frac{F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} < +\infty, \quad \alpha := 2\sqrt{2}c\beta^2\gamma < 1. \quad (17)$$

Take  $0 \leq \vartheta < 1$ ,  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1(\alpha + \alpha\vartheta + \vartheta) + \omega_2 < 1$ . Let  $a := (1 - \vartheta\omega_1 - \omega_2)$ ,  $b := (1 + \vartheta)\omega_1\beta$ ,  $\bar{a} := b + 2a(1 + \beta) - \sqrt{2}\gamma\beta bc$  and

$$r := \min \left\{ \kappa, \frac{\bar{a} - \sqrt{\bar{a}^2 - 4a(1 + \beta)(a - 2\sqrt{2}c\beta b\gamma)}}{2a\gamma(1 + \beta)} \right\}.$$

Then, the inexact Gauss-Newton like methods for solving (1), with initial point  $x_0 \in B(x_*, r) \setminus \{x_*\}$

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F'(x_k)^*F(x_k) + r_k, \quad k = 0, 1, \dots, \quad (18)$$

with the following conditions for the residual  $r_k$ , and the forcing term  $\theta_k$

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^*F(x_k)\|, \quad 0 \leq \theta_k \text{cond}(P_k F'(x_k)^*F'(x_k)) \leq \vartheta, \quad k = 0, 1, \dots,$$

where  $\{P_k\}$  is an invertible matrix sequence (preconditioners for the linear system in (18)) and  $B(x_k)$  is an invertible approximation of  $F'(x_k)^*F'(x_k)$  satisfying

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

is well defined, contained in  $B(x_*, r)$ , converges to  $x_*$  and there holds

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \frac{(1 + \vartheta)\omega_1\beta\gamma}{(1 - \gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} \|x_k - x_*\|^2 \\ &\quad + \left( \frac{(1 + \vartheta)\omega_1\sqrt{2}c\beta^2\gamma(2 - \gamma\|x_0 - x_*\|)}{(1 - \gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} + \omega_1\vartheta + \omega_2 \right) \|x_k - x_*\|, \end{aligned}$$

for all  $k = 0, 1, \dots$ .

We need the following result to prove the above theorem.

**Lemma 19.** Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $F : \Omega \rightarrow \mathbb{Y}$  an analytic function. Suppose that  $x_* \in \Omega$  and  $B(x_*, 1/\gamma) \subset \Omega$ , where  $\gamma$  is defined in (17). Then, for all  $x \in B(x_*, 1/\gamma)$  there holds

$$\|F''(x)\| \leq 2\gamma/(1 - \gamma\|x - x_*\|)^3.$$

*Proof.* See the proof of the Lemma 21 of [16]. □



The next result gives a condition that is easier to check than condition (2), whenever the functions under consideration are twice continuously differentiable.

**Lemma 20.** *Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $x_* \in \Omega$  and  $F : \Omega \rightarrow \mathbb{Y}$  be twice continuously on  $\Omega$ . If there exists a  $f : [0, R) \rightarrow \mathbb{R}$  twice continuously differentiable such that*

$$\|F''(x)\| \leq f''(\|x - x_*\|), \quad (19)$$

*for all  $x \in \Omega$  such that  $\|x - x_*\| < R$ . Then  $F$  and  $f$  satisfy (2).*

*Proof.* See the proof of the Lemma 22 of [16]. □

**[Proof of Theorem 18].** Consider the real function  $f : [0, 1/\gamma) \rightarrow \mathbb{R}$  defined by

$$f(t) = \frac{t}{1 - \gamma t} - 2t.$$

It is straightforward to show that  $f$  is analytic and that

$$f(0) = 0, \quad f'(t) = 1/(1 - \gamma t)^2 - 2, \quad f'(0) = -1, \quad f''(t) = (2\gamma)/(1 - \gamma t)^3, \quad f^n(0) = n! \gamma^{n-1},$$

for  $n \geq 2$ . It follows from the last equalities that  $f$  satisfies **h1** and **h2**. Since  $2\sqrt{2}c\beta^2\gamma < 1$  the condition **h3** also holds. Now, as  $f''(t) = (2\gamma)/(1 - \gamma t)^3$  combining Lemmas 20, 19 we conclude that  $F$  and  $f$  satisfy (2) with  $R = 1/\gamma$ . In this case, it is easy to see that constants  $\nu$  and  $\rho$  as defined in Theorem 7, satisfy

$$0 < \rho = \frac{\bar{a} - \sqrt{\bar{a}^2 - 4a(1 + \beta)(a - 2\sqrt{2}c\beta b\gamma)}}{2a\gamma(1 + \beta)} < \nu = ((1 + \beta) - \sqrt{\beta(1 + \beta)})/(\gamma(1 + \beta)) < 1/\gamma,$$

and as a consequence,  $0 < r = \min\{\kappa, \rho\}$ . Therefore, as  $F$ ,  $\sigma$ ,  $f$  and  $x_*$  satisfy all hypotheses of Theorem 7, taking  $x_0 \in B(x_*, r) \setminus \{x_*\}$ , the statements of the theorem follow from Theorem 7. □

For the case  $\vartheta = 0$ , the Theorem 18 becomes:

**Corollary 21.** *Let  $\Omega \subseteq \mathbb{X}$  be an open set,  $F : \Omega \rightarrow \mathbb{Y}$  a continuously differentiable function. Let  $x_* \in \Omega$ ,  $R > 0$  and*

$$c := \|F(x_*)\|, \quad \beta := \|F'(x_*)^\dagger\|, \quad \kappa := \sup\{t \in [0, R) : B(x_*, t) \subset \Omega\}.$$

*Suppose that  $F'(x_*)^*F(x_*) = 0$ ,  $F'(x_*)$  is injective and*

$$\gamma := \sup_{n \geq 1} \left\| \frac{F^{(n)}(x_*)}{n!} \right\|^{1/(n-1)} < +\infty, \quad \alpha := 2\sqrt{2}c\beta^2\gamma < 1.$$

Take  $0 \leq \omega_2 < \omega_1$  such that  $\omega_1\alpha + \omega_2 < 1$ . Let  $\bar{a} := \omega_1\beta + 2(1 - \omega_2)(1 + \beta) - \sqrt{2}\gamma\beta^2\omega_1c$  and

$$r := \min \left\{ \kappa, \frac{\bar{a} - \sqrt{\bar{a}^2 - 4(1 - \omega_2)(1 + \beta)(1 - \omega_2 - 2\sqrt{2}c\beta^2\omega_1\gamma)}}{2(1 - \omega_2)\gamma(1 + \beta)} \right\}.$$

Then, the Gauss-Newton's like method for solving (1), with initial point  $x_0 \in B(x_*, r) \setminus \{x_*\}$

$$x_{k+1} = x_k + S_k, \quad B(x_k)S_k = -F'(x_k)^*F(x_k), \quad k = 0, 1, \dots,$$

where  $B(x_k)$  is an invertible approximation of  $F'(x_k)^*F'(x_k)$  satisfying

$$\|B(x_k)^{-1}F'(x_k)^*F'(x_k)\| \leq \omega_1, \quad \|B(x_k)^{-1}F'(x_k)^*F'(x_k) - I\| \leq \omega_2, \quad k = 0, 1, \dots,$$

is well defined, contained in  $B(x_*, r)$ , converges to  $x_*$  and there holds

$$\begin{aligned} \|x_{k+1} - x_*\| &\leq \frac{\omega_1\beta\gamma}{(1 - \gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} \|x_k - x_*\|^2 \\ &\quad + \left( \frac{\omega_1\sqrt{2}c\beta^2\gamma(2 - \gamma\|x_0 - x_*\|)}{(1 - \gamma\|x_0 - x_*\|)^2 - \beta\gamma(2\|x_0 - x_*\| - \gamma\|x_0 - x_*\|^2)} + \omega_2 \right) \|x_k - x_*\|, \end{aligned}$$

for all  $k = 0, 1, \dots$ .

Note that letting  $c = 0$  in the above corollary, we obtain the Example 1 of [8].

## 4 Final remark

The Theorem 7 gives an estimate of the convergence radius for inexact Gauss-Newton like methods. In particular, for  $\vartheta = \omega_1 = 0$  and  $\omega_2 = 1$  is shown in Ferreira et al. [16], that  $r$  is the best possible convergence radius.

Another detail is that, as pointed out by Morini in [23] if preconditioning  $P_k$ , satisfying

$$\|P_k r_k\| \leq \theta_k \|P_k F'(x_k)^*F(x_k)\|, \quad (20)$$

for some forcing sequence  $\{\theta_k\}$ , is applied to finding the inexact Gauss-Newton steep, then the inverse proportionality between each forcing term  $\theta_k$  and  $\text{cond}(P_k F'(x_k)^*F(x_k))$  stated in the following assumption:

$$0 < \theta_k \text{cond}(P_k F'(x_k)^*F(x_k)) \leq \vartheta, \quad k = 0, 1, \dots, \quad (21)$$

is sufficient to guarantee convergence, and may be overly restrictive to bound the sequence  $\{\theta_k\}$ , always such that the matrices  $P_k F'(x_k)^*F(x_k)$ , for  $k = 0, 1, \dots$ , are badly conditioned. Moreover,  $\theta_k$  does not depend on  $\text{cond}(F'(x_k)^*F(x_k))$  but only on the  $\text{cond}(P_k F'(x_k)^*F(x_k))$  and a suitable choice of scaling matrix  $P_k$  leads to a relaxation of the forcing terms.

## References

- [1] R. L. Adler, J. P. Dedieu, J. Y. Margulies, M. Martens, and M. Shub. Newton's method on Riemannian manifolds and a geometric model for the human spine. *IMA J. Numer. Anal.*, 22(3):359–390, 2002.
- [2] F. Alvarez, J. Bolte, and J. Munier. A unifying local convergence result for Newton's method in Riemannian manifolds. *Found. Comput. Math.*, 8(2):197–226, 2008.
- [3] I. K. Argyros and S. Hilout. Improved generalized differentiability conditions for Newton-like methods. *J. Complexity*, In Press, Corrected Proof:–, 2010.
- [4] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and real computation*. Springer-Verlag, New York, 1998. With a foreword by Richard M. Karp.
- [5] J. Chen. The convergence analysis of inexact Gauss-Newton methods for nonlinear problems. *Comput. Optim. Appl.*, 40(1):97–118, 2008.
- [6] J. Chen and W. Li. Convergence of Gauss-Newton's method and uniqueness of the solution. *Appl. Math. Comput.*, 170(1):686–705, 2005.
- [7] J. Chen and W. Li. Convergence behaviour of inexact Newton methods under weak Lipschitz condition. *J. Comput. Appl. Math.*, 191(1):143–164, 2006.
- [8] J. Chen and W. Li. Local convergence results of Gauss-Newton's like method in weak conditions. *J. Math. Anal. Appl.*, 324(2):1381 – 1394, 2006.
- [9] J. P. Dedieu and M. H. Kim. Newton's method for analytic systems of equations with constant rank derivatives. *J. Complexity*, 18(1):187–209, 2002.
- [10] J. P. Dedieu, P. Priouret, and G. Malajovich. Newton's method on Riemannian manifolds: covariant alpha theory. *IMA J. Numer. Anal.*, 23(3):395–419, 2003.
- [11] J. P. Dedieu and M. Shub. Newton's method for overdetermined systems of equations. *Math. Comp.*, 69(231):1099–1115, 2000.
- [12] R. S. Dembo, S. C. Eisenstat, and T. Steihaug. Inexact Newton methods. *SIAM J. Numer. Anal.*, 19(2):400–408, 1982.
- [13] J. E. Dennis, Jr. and R. B. Schnabel. *Numerical methods for unconstrained optimization and nonlinear equations*, volume 16 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1996. Corrected reprint of the 1983 original.
- [14] O. P. Ferreira. Local convergence of Newton's method in Banach space from the viewpoint of the majorant principle. *IMA J. Numer. Anal.*, 29(3):746–759, 2009.

- [15] O. P. Ferreira and M. L. N. Goncalves. Local convergence analysis of inexact Newton-like methods under majorant condition. *Comput. Optim. Appl.*, Article in Press:1–21, 2009.
- [16] O. P. Ferreira, M. L. N. Goncalves, and P. R. Oliveira. Local convergence analysis of gauss-newton’s method under majorant condition. To appear in *J. Complexity*, 1–20, 2010.
- [17] O. P. Ferreira and B. F. Svaiter. Kantorovich’s majorants principle for Newton’s method. *Comput. Optim. Appl.*, 42(2):213–229, 2009.
- [18] J.-B. Hiriart-Urruty and C. Lemaréchal. *Convex analysis and minimization algorithms. I*, volume 305 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1993. Fundamentals.
- [19] L. V. Kantorovich. The principle of the majorant and Newton’s method. *Doklady Akad. Nauk SSSR (N.S.)*, 76:17–20, 1951.
- [20] C. Li, N. Hu, and J. Wang. Convergence behavior of Gauss-Newton’s method and extensions of the smale point estimate theory. *J. Complexity*, In Press, Corrected Proof:–, 2010.
- [21] C. Li, W. H. Zhang, and X. Q. Jin. Convergence and uniqueness properties of Gauss-Newton’s method. *Comput. Math. Appl.*, 47(6-7):1057–1067, 2004.
- [22] J. M. Martinez and L. Qi. Inexact Newton methods for solving nonsmooth equations. *J. Comput. Appl. Math.*, 60:127–145, 1999.
- [23] B. Morini. Convergence behaviour of inexact Newton methods. *Math. Comp.*, 68:1605–1613, 1999.
- [24] P. D. Proinov. General local convergence theory for a class of iterative processes and its applications to Newton’s method. *J. Complexity*, 25(1):38 – 62, 2009.
- [25] P. D. Proinov. New general convergence theory for iterative processes and its applications to Newton-Kantorovich type theorems. *J. Complexity*, 26(1):3 – 42, 2010.
- [26] S. Smale. Newton’s method estimates from data at one point. In *The merging of disciplines: new directions in pure, applied, and computational mathematics (Laramie, Wyo., 1985)*, pages 185–196. Springer, New York, 1986.
- [27] G. W. Stewart. On the continuity of the generalized inverse. *SIAM J. Appl. Math.*, 17:33–45, 1969.
- [28] X. Wang. Convergence of Newton’s method and uniqueness of the solution of equations in Banach space. *IMA J. Numer. Anal.*, 20(1):123–134, 2000.
- [29] P. A. Wedin. Perturbation theory for pseudo-inverses. *Nordisk Tidskr. Informationsbehandling (BIT)*, 13:217–232, 1973.